Eötvös Loránd Tudományegyetem

Természettudományi Kar

Hegedűs Dávid

Nemlineáris egyenletrendszerek numerikus megoldása

BSc Szakdolgozat

Témavezető:

Dr. Gergó Lajos

Numerikus Analízis Tanszék

Budapest, 2014

Köszönetnyilvánítás

Szeretném megköszönni Dr. Gergó Lajos tanár úrnak, hogy elvállalta a konzulensi feladatkört, valamint ötleteivel, tanácsaival nagy mértékben segítette a munkámat. Köszönettel tartozom a családomnak és a barátaimnak, akik kitartóan támogattak az egyetemi éveim alatt.

Tartalomjegyzék

Bevezetés

Az alkalmazott matematika számos területén (pl. fizika, robotika, irányítástechnika) sokszor találkozhatunk olyan problémákkal, amelyekben több változó szerepel és közöttük több összefüggés is fennáll. Ha a változók legfeljebb az első hatványon szerepelnek, akkor lineáris egyenletrendszert kapunk, azonban a természettudományos, technológiai és gazdasági folyamatok matematikai modelljei sokszor nemlineárisak, tehát magasabb hatványkitevők is lehetnek. Sok esetben arra vagyunk kíváncsiak, hogy milyen feltételek mellett lesz a megadott függvényünk nulla. Ekkor az $f(x) = 0$ nemlineáris egyenletrendszer megoldásait közelíthetjük különböző numerikus módszerekkel. A szakdolgozatomban ezt a feladatot a Newton-módszerrel és annak változataival fogom megoldani. Ezen iterációs módszereknek az az előnye, hogy gyorsabban mint lineárisan konvergálnak, viszont általában nem minden kezdővektorra. Előfordulhat, hogy már a futtatáshoz szükséges információk megszerzése is nehézkes (kezdőpont, derivált), ekkor csillapítással vagy a Broyden-módszerrel próbálkozhatunk. Magasabb dimenzióban fontos a konvergencia globalizálása akár a futási idő kárára is, ezt fogjuk látni a folytatás módszerénél. A dolgozat végén alkalmazásokat tekinthet meg az Olvasó, melyekhez MATLAB segítségével készítettem programokat. A programok M-fájljai a CD-mellékleten találhatók, az utolsó fejezetben segítséget nyújtok a használatukhoz.

1. fejezet

Nemlineáris egyenletek numerikus megoldása

1.1. A Newton-módszer

Legyen $f : \mathbb{R} \to \mathbb{R}$ függvény, amely folytonos az $[a, b]$ intervallumon, $a < b$. Ekkor ha f előjelet vált a és b között, akkor az

$$f(x) = 0 \tag{1.1}$$

egyenletnek van gyöke $[a, b]$-ben. A Newton-módszer egy olyan iterációs eljárás, melynek során egy x_0 kezdőpontból szeretnénk közelíteni (1.1) egyenlet egyik gyökét. Ehhez keressünk egy δx javítást, melyre $x_0 + \delta x$ már gyök lesz. Tegyük fel, hogy $f(x)$ kétszer folytonosan differenciálható és tekintsük a Taylor-sorát:

$$0 = f(x_0 + \delta x) = f(x_0) + f'(x_0)\delta x + O(\delta x^2).$$

Legyen $x_1 := x_0 + \delta x$ és hanyagoljuk el a másodrendű tagot. Ekkor a

$$0 = f(x_0) + f'(x_0)(x_1 - x_0) \tag{1.2}$$

egyenlet adódik. Feltéve, hogy $f'(x_0) \neq 0$, x_1 átrendezéssel kiszámítható:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Értelemszerűen folytatva, adott x_0 esetén a következő iterációs eljárást kapjuk:

$$x_{k+1} = x_k - (f'(x_k))^{-1} f(x_k), \quad k = 0, 1, \ldots. \tag{1.3}$$

Egy dimenziós esetben jól látható a Newton módszer mértani jelentése. Húzzunk a függvény $f(x_0)$ pontjához érintőt. Az érintő egyenlete a következő:

$$y = f(x_0) + f'(x_0)(x - x_0).$$

Bizonyos feltételek mellett azt várjuk, hogy ennek az egyenesnek a gyöke az f gyökének egy jobb közelítése lesz. Ha az érintő zérushelyét x_1 jelöli, akkor

$$0 = f(x_0) + f'(x_0)(x_1 - x_0),$$

ez pedig megegyezik az (1.2) egyenlettel, tehát az eljárást folytatva következik az (1.3) iteráció.

1.1.1. Definíció. *Tegyük fel, hogy egy iterációs eljárás az $x_k \in \mathbb{R}^n$ sorozatot állítja elő, ez a sorozat konvergál a $\|.\|$ normában, $\lim x_k = x_*$, $x_0 \neq x_*$. Azt mondjuk, hogy az x_k sorozat x_0-ból kiindulva, (legalább) κ-rendben konvergál, ha $\kappa \geq 1$, és van olyan nemnegatív C konstans, hogy érvényes*

$$\|x_{k+1} - x_*\| \leqslant C\|x_k - x_*\|^\kappa, \quad k = 0, 1, \ldots \quad .$$

A $\kappa > 1$ renddel való konvergenciát szuperlineáris konvergenciának nevezzük.

1.1.2. Definíció. *Egy iterációs eljárásról akkor mondjuk, hogy lokálisan konvergál az x_* megoldáshoz, ha van olyan pozitív δ sugarú $G(x_*, \delta)$ gömb, amely a következő tulajdonságokkal rendelkezik: tetszőleges $x_0 \in G(x_*, \delta)$ esetén az iteráció által előállított x_k sorozat konvergál x_*-hoz; $G(x_*, \delta)$-n kívül viszont van olyan x_0 vektor, amelyre ez nem igaz. $G(x_*, \delta)$ az x_* vonzási gömbje.*

1.1.3. Definíció. *Egy iterációs eljárásról akkor mondjuk, hogy globálisan konvergál az x_* megoldáshoz, ha tetszőleges x_0 kezdővektor esetén az iteráció által előállított x_k sorozat konvergál x_*-hoz.*

1.1.1. Konvergencia tétel

1.1.4. Tétel. (A Newton-módszer konvergenciája) *Legyen $I := (a, b)$, $f : I \to \mathbb{R}$ olyan függvény, hogy a deriváltja Lipschitz-folytonos I-ben L Lipschitz állandóval, azaz: $|f'(x) - f'(y)| \leqslant L|x - y| \quad \forall x, y \in I$;*

továbbá létezzék egy $\alpha > 0$, melyre $|f'(x)| \geq \frac{1}{\alpha}$, $\forall x \in I$.
Ha az $f(x) = 0$ egyenletnek I-ben van x_* gyöke, akkor létezik olyan $\delta > 0$, hogy $x_0 \in I$-ből és $|x_0 - x_*| < \delta$-ból következik, hogy:

1. *a* (1.3) *sorozat jól definiált;*

2. $\lim_{k\to\infty} x_k = x_*$*;*

3. $|x_{k+1} - x_*| \leqslant C|x_k - x_*|^2, \quad C := \frac{L\alpha}{2}$.

Bizonyítás. Kezdetben legyen $k = 0$, $f_k := f(x_k)$. A feltételből következik, hogy $f'_k := f'(x_k) \neq 0$. Végezzük el a következő átalakításokat:

$$x_{k+1}-x_* = x_k-x_*-(f'_k)^{-1}f_k = x_k-x_*-(f'_k)^{-1}(f_k-f_*) = -(f'_k)^{-1}(f_k-f_*-f'_k(x_k-x_*)),$$

ahol $f_* := f(x_*) = 0$. Legyen $r(x_k, x_*) := (f_k - f_* - f'_k(x_k - x_*)$, és külön becsüljük általános $x, y \in I$-re. A Newton-Leibniz formula felhasználásával az

$$r(x,y) = f(x) - f(y) - f'(x)(x-y) = \int_y^x \big(f'(z) - f'(x)\big)dz$$

azonosság adódik. A $z = y + t(x - y)$ helyettesítéssel következik, hogy

$$r(x,y) = (x-y)\int_0^1 \big(f'(y + t(x-y)) - f'(x)\big)dt \quad \big(t \in [0,1]\big).$$

Most használjuk ki, hogy f' Lipschitz-folytonos:

$$|r(x,y)| \leqslant L(x-y)^2 \int_0^1 (1-t)dt = \frac{L}{2}(x-y)^2.$$

Ezt a becslést, valamint az $|f'_k| \geq \frac{1}{\alpha}$ feltételt felhasználva kapjuk a következőt:

$$|x_{k+1} - x_*| \leqslant \frac{L\alpha}{2}|x_k - x_*|^2.$$

A konvergencia belátásához válasszunk olyan $|x_k - x_*| \leqslant \delta$-t, hogy $x_k \in I$ és

$$\frac{L\alpha|x_k - x_*|}{2} \leqslant q < 1$$

teljesüljön, hiszen ekkor igaz a következő:

$$|x_{k+1} - x_*| \leqslant q|x_k - x_*| \leqslant |x_k - x_*|, \quad k = 0, 1, \ldots \quad .$$

Alkalmas δ-val elérhetjük, hogy $x_{k+1} \in I$. Ilyenkor már x_{k+2} is definiálva van, és

$$|x_{k+2} - x_*| \leqslant \frac{L\alpha}{2}|x_{k+1} - x_*|^2 \leqslant q|x_{k+1} - x_*| \leqslant q^2|x_k - x_*| \leqslant |x_k - x_*| \leqslant \delta,$$

tehát $x_{k+2} \in I$. Folytatva az eljárást láthatjuk, hogy ez pont a konvergenciát jelenti. □

Az (1.1.4) tételben látjuk, hogy a Newton-módszer másodrendben (másnéven négyzetesen vagy kvadratikusan) konvergens. Továbbá az is kiderül, hogy ehhez x_0 elég közel kell legyen x_*-hoz, tehát (általában) csak lokálisan konvergens. A későbbiekben lesz még szó szuperlineárisan konvergens módszer „globalizálásáról".

1.2. A csillapított Newton-módszer

A Newton-módszer alkalmazása során különböző problémák merülhetnek fel. Például, ha a kezdőpontban a függvény deriváltja nulla, akkor nem tudjuk meghatározni a következő pontot. Ha pedig a választott kezdőpont nincs elég közel a gyökhöz, akkor előfordulhat, hogy az iteráció lassabban vagy egyáltalán nem fog konvergálni, esetleg végtelen ciklusba kerül.

A továbbiakban azt az esetet fogom megvizsgálni, amikor x_* többszörös gyök. Ekkor ugyanis $f'(x_*) = 0$, és az ilyen esetet eddig kizártuk a tárgyalásból. Legyen f elegendően sokszor differenciálható, és rendelkezzen t-szeres gyökkel x_*-ban, $t \geq 1$. Ekkor Taylor-sorfejtésének eleje a következő:

$$f(x) = (x - x_*)^t \left[\frac{1}{t!}f^{(t)}(x_*) + \frac{x - x_*}{(t+1)!}f^{(t+1)}(x_*) + \dots\right], \quad f^{(t)}(x_*) \neq 0.$$

Legyen $e_k := x_k - x_*$. Ekkor az alábbi egyenleteket kapjuk:

$$f(x_k) = \frac{(e_k)^t}{t!}\left[f^{(t)}(x_*) + \frac{e_k}{t+1}f^{(t+1)}(x_*) + O(e_k^2)\right],$$

$$f'(x_k) = \frac{(e_k)^{t-1}}{(t-1)!}\left[f^{(t)}(x_*) + \frac{e_k}{t}f^{(t+1)}(x_*) + O(e_k^2)\right].$$

Ezek felhasználásával az $e_{k+1} = e_k - \dfrac{f(x_k)}{f'(x^{(k)})}$-ból elegendően kicsi e_k-ra következik, hogy

$$e_{k+1} = e_k - \frac{e_k}{t}\left[f^{(t)}(x_*) + \frac{e_k}{t+1}f^{(t+1)}(x_*) + O(e_k^2)\right] *$$
$$* f^{(t)}(x_*)^{-1}\left[1 - \frac{e_k}{tf^{(t)}(x_*)}f^{(t+1)}(x_*) + O(e_k^2)\right],$$

,mert

$$\left[1+\frac{e_k}{tf^{(t)}(x_*)}f^{(t+1)}(x_*)+O(e_k^2)\right]^{-1}=\left[1-\frac{e_k}{tf^{(t)}(x_*)}f^{(t+1)}(x_*)+O(e_k^2)\right].$$

Tehát:

$$e_{k+1}=e_k-\frac{e_k}{tf^{(t)}(x_*)}\left[f^{(t)}(x_*)-\frac{e_k}{t(t+1)}f^{(t+1)}(x_*)+O(e_k^2)\right]=$$

$$=\left(1-\frac{1}{t}\right)e_k+e_k^2\frac{f^{(t+1)}(x_*)}{t^2(t+1)f^{(t)}(x_*)}+O(e_k^3).$$

Láthatjuk, hogy $t=1$-re négyzetes a konvergencia, azaz a klasszikus Newton-módszer hajtódik végre. Viszont ha $t>1$, akkor csak első rendben, vagyis lineárisan konvergens.
Ha t ismert, akkor a Newton eljárás következő módosítása biztosítja újra a négyzetes konvergenciát:

$$x_0 \text{ adott}, \quad x_{k+1}=x_k-t(f'(x_k))^{-1}f(x_k), \quad k=0,1,\ldots.$$

Tehát ebben az esetben az eredeti Newton-lépést hosszabbítani kellett ahhoz, hogy javuljon a konvergencia. Általános esetben viszont fordítva érdemes eljárni: $t<1$ esetén ugyanis a Newton-módszer lokális konvergenciáját globálissá terjeszthetjük ki, ugyanakkor a konvergencia sebessége csökken. Ezt nevezzük csillapított Newton-módszernek. A gyakorlatban ez az eljárás általában eléri azt a gyök-közelítést, amelyet a klasszikus módszernek adtunk meg, onnantól kezdve pedig $t=1$ teljesül.

1.3. A szelőmódszer

A Newton-módszer másik gondja a derivált kiszámítása, amely fokozottan jelentkezik egyenletrendszerek esetében. Ez nagy jelentőséggel bír, hiszen az alkalmazás során előfordulhat, hogy a deriváltfüggvény nem áll a rendelkezésünkre, nem tudjuk kiszámítani, esetleg f-re nem is létezik zárt képlet. A feladatunk tehát az, hogy kiküszöböljük az iterációs képletből a deriváltat, és helyettesítsük annak valamilyen közelítésével. Erre egy lehetőség a differenciahányadossal való helyettesítés:

$$f'(x_k)\approx\frac{f(x_k+h)-f(x_k)}{h}=:A_k$$

Így az iterációt a következő alakban írhatjuk fel:

$$x_0 \text{ adott}, \quad x_{k+1} = x_k - \frac{f(x_k)}{A_k}, \quad k = 0, 1, \ldots .$$

Amennyiben f kétszer folytonosan differenciálható (ezzel némileg megszorítva a konvergencia tétel feltételeit), akkor a differenciaképlet első renddel approximálja a deriváltat.

Legyen $M_2 := max_{[a,b]}|f''(x)|$. Ekkor tehát:

$$|f'(x_k) - A_k| \leqslant \frac{1}{2} M_2 |h|.$$

Belátható, hogy létezik olyan h-kiválasztás, hogy az iteráció konvergenciája továbbra is másodrendű marad, annak ellenére, hogy x_* ismeretlen.

Egy másik lehetőség, ha az $f'(x_k)$-t az $[x_{k-1}, x_k]f$ osztott differenciával helyettesítjük. Ennek előnye, hogy ebben az esetben nincs szükség külön f-érték kiszámítására:

$$x_0 \text{ és } x_1 \neq x_0 \text{ adott}, \quad x_{k+1} = x_k - \frac{f(x_k)}{[x_{k-1}, x_k]f}, \quad k = 0, 1, \ldots .$$

1.3.1. Tétel. (A szelőmódszer konvergenciája) *Teljesüljenek a Newton-módszer konvergencia tételének feltételei, legyen $f \in C^2[a,b]$. Ekkor a szelőmódszer vagy véges k-ra megáll az $f(x) = 0$ megoldásán vagy (lokálisan) konvergens $\frac{1}{2}(1+\sqrt{5})$ renddel.*

1.3.2. Megjegyzés. Vegyük észre, hogy a Newton-módszer egy függvényértéket és egy deriváltértéket követel meg egy iterációs lépésben. Ugyanakkor a szelőmódszer esetében csak egyetlen függvénykiértékelésre van szükségünk. Feltehetjük, hogy ugyanannyi időbe kerül kiszámítani f-et és f'-t. Ekkor egy Newton-lépés alatt két lépést lehet végrehajtani a szelőmódszerrel, tehát az utóbbi egyértelműen gyorsabb eljárás.

1.4. Monoton konvergencia

Igazolható, hogy létezik olyan függvényosztály (a konvex vagy konkáv függvények osztálya), amelyre a Newton-módszer monoton és globális konvergenciája egyaránt teljesül. Ekkor nincs szükségünk arra, hogy x_0 elég jó közelítés legyen, és arra sem, hogy csillapított módszert használjunk.

1.4.1. Tétel. *Legyen* $f \in C^2[a,b]$, $f'(x) \neq 0$ *és* $f''(x) \neq 0$ *az* $[a,b]$ *intervallumon, amely végtelen is lehet. Létezzen* $x_* \in (a,b)$, *melyre* $f(x_*) = 0$. *Ekkor az* (1.3)*-ban definiált iterációban* $x_k \to x_*$ *monoton módon, ha* $f(x_0)f''(x_0) > 0$ *és* $x_0 \in [a,b]$.

1.4.2. Megjegyzés. A tétel lényeges feltételei az $f'(x) \neq 0$ és az $f''(x) \neq 0$. Ugyanis, ha megengednénk, hogy az $f'(x) = 0$ legyen valamely x-ekre, akkor ezeken a helyeken a függvényhez húzott érintők párhuzamosak lennének az x-tengellyel, tehát az iteráció nem tudna továbblépni. Ha pedig az $f''(x) \neq 0$ feltétel nem teljesül (azaz sérül a konvexitás vagy a konkávitás), akkor $\{x_k\}$ sorozat bonyolultan viselkedhet, esetleg végtelen ciklusba kerülhet.
Kevésbé fontos az $f(x_0)f''(x_0) > 0$ feltétel, mivel a nem teljesülése esetén is monoton lesz a konvergencia x_1-től kezdve.
Lássunk erre egy példát: legyen $f''(x) > 0$ és $f(x_0) < 0$. Ekkor $f'(x) > 0$ és igazak az alábbiak:

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0),$$

$$f(x_1) \geq f(x_0) + f'(x_0)(x_1 - x_0) = 0.$$

Tehát $x_0 < x_* \leqslant x_1$, azaz x_1-re már teljesül az $f(x_1)f''(x_1) > 0$ feltétel.

2. fejezet

Nemlineáris egyenletrendszerek numerikus megoldása

2.1. A többdimenziós Newton-módszer

Nemlineáris egyenletrendszerek megoldása gyakran bizonytalan kimenetelű, fordulatokban gazdag vállalkozás. Megoldásukra szinte kizárólag a Newton-módszer változatait használjuk, mivel már az első deriváltak meghatározása se mindig könnyű feladat, továbbá a magasabbrendű módszerek még az egydimenziós esetben sem feltétlenül hatékonyabbak.

Ebben a fejezetben a felső indexek nem hatványkitevőket, hanem vektorok komponenseit fogják jelölni. Legyen $f : \mathbb{R}^n \to \mathbb{R}^n$.

A feladat a következő: keressük az

$$f(x) = \begin{bmatrix} f_1(x^1, \ldots, x^n) \\ \vdots \\ f_n(x^1, \ldots, x^n) \end{bmatrix} = 0 \tag{2.1}$$

rendszernek egy megoldását. Tegyük fel, hogy f differenciálható, x_* az f egy megoldása és x_0 az x_*-nak egy közelítése. Ekkor a másodrendű tagok elhanyagolásával kapott Taylor-polinom a következő:

$$0 = f(x_*) \approx f(x_0) + Df(x_0)(x_* - x_0),$$

,ahol

$$Df(x_0) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x^1} & \cdots & \dfrac{\partial f_1}{\partial x^n} \\ \vdots & & \vdots \\ \dfrac{\partial f_n}{\partial x^1} & \cdots & \dfrac{\partial f_n}{\partial x^n} \end{bmatrix}_{x=x_0}, \qquad x_* - x_0 = \begin{bmatrix} x_*^1 - x_0^1 \\ \vdots \\ x_*^n - x_0^n \end{bmatrix}.$$

Ha a $Df(x_0)$ Jacobi-mátrix reguláris, akkor az

$$f(x_0) + Df(x_0)(x_1 - x_0) = 0$$

egyenletből x_1-et kifejezve a következőt kapjuk:

$$x_1 = x_0 - (Df(x_0))^{-1} f(x_0),$$

,ahol (bizonyos feltételek teljesülése mellett) x_1 az x_0 javítását adja, tehát jobb közelítése lesz x_*-nak. Így a (2.1) egyenletrendszer megoldására szolgáló iterációs eljárás:

$$x_0 \text{ adott}, \quad x_{k+1} = x_k - (Df(x_k))^{-1} f(x_k), \quad k = 0, 1, 2, \ldots.$$

2.1.1. Megjegyzés. Ha a $Df(x_0)$ Jacobi-mátrix szinguláris, akkor más x_0 kezdővektort kell megadnunk.

2.1.2. Megjegyzés. Valójában az iteráció minden lépésében egy lineáris egyenletrendszert oldunk meg, ezért célszerű a következő alakban megadni:

$$x_0 \text{ adott}, \quad Df(x_k)\delta x_k = -f(x_k), \quad x_{k+1} = x_k + \delta x_k, \quad k = 0, 1, 2, \ldots.$$

Láthatjuk, hogy ilyenkor az inverz mátrixot nem számítjuk ki, helyette alkalmazzuk például a Gauss-elimináció megfelelő változatát.

A Newton-módszer minden lépésében meg kell határoznunk a Jacobi-mátrixot, vagyis n^2 függvényértéket. Ez problémát jelenthet az alkalmazások során, ezért (az egydimenziós eset mintájára) differenciaképletekkel közelíthetjük a deriváltakat. Többdimenziós esetben diszkretizált Newton-módszerről beszélünk, ha minden k. lépésben az $Df(x_k)$ Jacobi-mátrix elemeit az alábbi elsőrendű közelítéssel számoljuk ki:

$$\frac{\partial f_i}{\partial x^j} \approx \frac{f_i(x_k + h_{ij}e_j) - f_i(x_k)}{h_{ij}}, \quad i, j = 1, \ldots, n, \tag{2.2}$$

, ahol e_j a j. koordináta-egységvektor és h_{ij} alkalmas nemnulla lépéstávolság. Könnyen láthatjuk, hogy ilyenkor $2n^2$ függvényérték meghatározására van szükségünk, tehát romlik a futási idő. Gyakori eset azonban, hogy a nemlineáris egyenletrendszerekhez tartozó Jacobi-mátrix ritkamátrix, ekkor pedig a közelítés műveletigénye jelentős mértékben csökkenthető. Például ha a mátrix tridiagonális, akkor csupán $3n - 2$ nemzérus elem van, tehát ennyi függvényérték kiszámítása elegendő.

2.1.1. Konvergencia tételek

A következőkben belátom, hogy a többdimenziós Newton-módszer kvadratikusan konvergens. Ennek a speciális esetét már láttuk az első fejezetben. Mindenekelőtt következzen két definíció és egy lemma, melyre szükség lesz a bizonyítás során.

2.1.3. Definíció. *Legyen $f : \mathbb{R}^n \to \mathbb{R}^n$, és tekintsük az $f(x) = 0$ egyenletrendszert. Az f függvényről azt mondjuk, hogy differenciálható egy $x_0 \in \mathbb{R}^n$ pontban, ha létezik $A : \mathbb{R}^n \to \mathbb{R}^n$ lineáris operátor, melyre*

$$\lim_{x \to x_0} \frac{\|f(x) - f(x_0) - A(x - x_0)\|}{\|x - x_0\|} = 0,$$

,ahol $\|.\|$ norma $\mathbb{R}^n$-en. Esetünkben A megegyezik a $Df(x_0)$ Jacobi-mátrixszal.

2.1.4. Definíció. *Legyen $C \subseteq \mathbb{R}^n$ tartomány. Azt mondjuk, hogy C konvex, ha minden $x, y \in C$-re és $t \in [0,1]$-re az $[x, y] := \{z = tx + (1-t)y\}$ halmaz is eleme C-nek.(Tehát x és y összekötő szakasza benne van C-ben.)*

2.1.5. Lemma. *Legyen $C_0 \subseteq \mathbb{R}^n$ konvex tartomány, $f : \mathbb{R}^n \to \mathbb{R}^n$. Ha f differenciálható C_0-on és létezik γ, melyre*

$$\|Df(x) - Df(y)\| \leqslant \gamma \|x - y\| \quad (\forall x, y \in C_0)$$

teljesül, akkor minden $x, y \in C_0$-ra igaz a következő becslés:

$$\|f(x) - f(y) - Df(y)(x - y)\| \leqslant \frac{\gamma}{2} \|x - y\|^2.$$

Bizonyítás. Definiáljuk a $\varphi : [0,1] \to \mathbb{R}^n$ függvényt a következőképpen:

$$\varphi(t) := f(y + t(x - y))$$

Ekkor φ differenciálható minden $t \in [0,1]$ helyen $(x, y \in C_0)$. Most határozzuk meg φ deriváltját a láncszabály alkalmazásával:

$$\varphi'(t) = Df(y + t(x-y))(x-y).$$

A lemma feltételét felhasználva a következő egyenlőtlenség írhaó fel:

$$\begin{aligned}\|\varphi'(t) - \varphi'(0)\| &= \|(Df(y + t(x-y)) - Df(y))(x-y)\| \leqslant \\ &\leqslant \|Df(y + t(x-y)) - Df(y)\|\|x-y\| \leqslant \\ &\leqslant \gamma t\|x-y\|^2 \qquad\qquad (0 \leqslant t \leqslant 1).\end{aligned}$$

Vezessünk be egy újabb jelölést! Legyen:

$$\begin{aligned}\Delta &:= f(x) - f(y) - Df(y)(x-y) = \varphi(1) - \varphi(0) - \varphi'(0) = \\ &= \int_0^1 (\varphi'(t) - \varphi'(0))dt.\end{aligned}$$

Mindezek felhasználásával könnyen adódik a lemma állítása:

$$\begin{aligned}\|\Delta\| = \|\int_0^1 (\varphi'(t) - \varphi'(0))dt\| &\leqslant \int_0^1 \|\varphi'(t) - \varphi'(0)\|dt \leqslant \\ &\leqslant \gamma\|x-y\|^2 \int_0^1 t\,dt = \frac{\gamma}{2}\|x-y\|^2.\end{aligned}$$

□

2.1.6. Tétel. *Legyen $C \subseteq \mathbb{R}^n$ nyílt halmaz, $C_0 \subseteq C$ egy konvex halmaz, valamint $f : C \to \mathbb{R}^n$ függvény, amely differenciálható minden $x \in C_0$ pontban és folytonos C-n. Legyen $x_0 \in C_0$, továbbá $r, \alpha, \beta, \gamma, h$ pozitív konstansok és teljesüljenek az alábbiak:*

$$\begin{aligned}S_r(x_0) &:= \{x : \|x - x_0\| < r\} \subseteq C_0, \\ h &:= \frac{\alpha\beta\gamma}{2} < 1, \\ r &:= \frac{\alpha}{1-h},\end{aligned}$$

és $f(x)$-re legyenek igazak a következő tulajdonságok:

a) $\|Df(x) - Df(y)\| \leqslant \gamma\|x-y\| \quad (\forall x, y \in C_0)$

b) $Df(x)^{-1}$ *létezik és teljesül, hogy* $\|Df(x)^{-1}\| \leqslant \beta, \quad \forall x \in C_0$*-ra* *c)* $\|Df(x_0)^{-1}f(x_0)\| \leqslant \alpha$.

Ekkor

1. *az alábbi sorozat jól definiált:*

$$x_0 \text{ adott}, \quad x_{k+1} = x_k - Df(x_k)^{-1}f(x_k), \quad k = 0, 1, \ldots$$

   és $x_k \in S_r(x_0)$ *minden* $k \geq 0$*-ra,*

2. $\lim_{k\to\infty} x_k = x_*$, $x_* \in S_r(x_0)$ *és* $f(x_*) = 0$,

3. *minden* $k \geq 0$*-ra*

$$\|x_k - x_*\| \leqslant \alpha \frac{h^{2^k-1}}{1-h^{2^k}}.$$

Mivel $0 < h < 1$*, ezért a Newton-módszer legalább másodrendben konvergens.*

Bizonyítás. (1.): A (b) feltétel szerint minden $x \in C_0$-ra létezik a $Df(x)^{-1}$, ezért az x_{k+1} sorozat jól definiált, ha $x_k \in S_r(x_0)$ minden $k \geq 0$-ra. Ez teljesül $k = 0$-ra és (c)-szerint $k = 1$-re. Teljes indukciót alkalmazunk: tegyük fel, hogy k-ig igaz az állítás, tehát $x_i \in S_r(x_0), \forall i = 2, \ldots, k$. Be kell látnunk, hogy $x_{k+1} \in S_r(x_0)$, ehhez tekintsük (b)-t és a következő egyenlőtlenséget:

$$\begin{aligned}\|x_{k+1} - x_k\| = \|Df(x_k)^{-1}f(x_k)\| \leqslant \beta\|f(x_k)\| = \\ = \beta\|f(x_k) - f(x_{k-1}) - Df(x_{k-1})(x_k - x_{k-1})\|,\end{aligned}$$

,ahol definíció szerint

$$f(x_{k-1}) + Df(x_{k-1})(x_k - x_{k-1}) = 0. \tag{2.3}$$

Most használjuk fel a (2.1.5)-ös lemmát:

$$\|x_{k+1} - x_k\| \leqslant \frac{\beta\gamma}{2}\|x_k - x_{k-1}\|^2.$$

Ezután vizsgáljuk meg a következő becslést:

$$\|x_{k+1} - x_k\| \leqslant \alpha h^{2^k-1}. \tag{2.4}$$

Jól látható, hogy a (c) tulajdonság miatt ez teljesül $k = 0$-ra. Ismét teljes indukciót alkalmazunk, feltesszük, hogy k-ig igaz és belátjuk, hogy $k + 1$-re is igaz marad:

$$\|x_{k+1} - x_k\| \leqslant \frac{\beta\gamma}{2}\|x_k - x_{k-1}\|^2 \leqslant \frac{\beta\gamma}{2}\alpha^2 h^{2^k-2} = \alpha h^{2^k-1}.$$

Felhasználva az előbbi becslést azonnal adódik az (1.) állítás:

$$\|x_{k+1}-x_0\| \leqslant \|x_{k+1}-x_k\| + \|x_k - x_{k-1}\| + \ldots + \|x_1 - x_0\| \leqslant$$
$$\leqslant \alpha(1+h+h^3+h^7+\ldots+h^{2^k-1}) < \alpha\frac{1}{1-h} = r,$$

ami azt jelenti, hogy $x_{k+1} \in S_r(x_0)$.
(2.): A (2.4) becslést felhasználva belátjuk, hogy $\{x_k\}$ egy Cauchy sorozat, tehát hogy $\forall\varepsilon > 0$-hoz $\exists n_\varepsilon$, hogy $\forall m \geq n > n_\varepsilon$-ra $\|x_m - x_n\| < \varepsilon$:

$$\|x_{m+1}-x_n\| \leqslant \|x_{m+1}-x_m\| + \|x_m - x_{m-1}\| + \ldots + \|x_{n+1} - x_n\| \leqslant \tag{2.5}$$
$$\leqslant \alpha h^{2^n-1}(1+h^{2^n}+(h^{2^n})^2+\ldots+(h^{2^n})^{2(m-n)-1}) < \frac{\alpha h^{2^n-1}}{1-h^{2^n}} < \varepsilon,$$

,mivel $h \in (0,1)$. Vegyük észre, hogy ez éppen a konvergenciát jelenti, vagyis létezik x_*, hogy

$$\lim_{k\to\infty} x_k = x_* \in \overline{S_r(x_0)},$$

,ahol a tartalmazás (1.)-ből következik.
Még azt kell megmutatni, hogy x_* az f zérushelye $\overline{S_r(x_0)}$-ban. Mivel $x_k \in S_r(x_0)$ minden $k \geq 0$-ra, valamint teljesül (a), ezért

$$\|Df(x_k) - Df(x_0)\| \leqslant \gamma\|x_k - x_0\| < \gamma r$$

teljesül. Definiáljuk K-t a következőképpen:

$$\|Df(x_k)\| \leqslant \gamma r + \|Df(x_0)\| =: K. \tag{2.6}$$

Átrendezve a (2.3) egyenletet azt kapjuk, hogy

$$f(x_k) = -Df(x_k)(x_{k+1} - x_k).$$

Ebből és (2.6)-ból mindkét oldal normáját véve

$$\|f(x_k)\| \leqslant K\|x_{k+1} - x_k\| \to 0 \quad (k \to \infty)$$

adódik, tehát:

$$\lim_{k\to\infty} \|f(x_k)\| = 0.$$

Végezetül mivel f folytonos x_*-ban, ezért

$$\lim_{k\to\infty} \|f(x_k)\| = \|f(x_*)\| = 0,$$

vagyis x_* az f zérushelye.
(3.): Ha (2.5)-ben $m \to \infty$, akkor az állítás azonnal látható:

$$\lim_{m\to\infty} \|x_m - x_n\| = \|x_* - x_n\| \leqslant \frac{\alpha h^{2^n-1}}{1-h^{2^n}}.$$

□

A feltételeket némileg erősítve igazolható, hogy a x_* az egyetlen zérushelye f-nek $S_r(x_0)$-ban. Erről szól az alábbi tétel:

2.1.7. Tétel. (Newton-Kantorovich) *Legyen $C \subseteq \mathbb{R}^n$ nyílt halmaz, $C_0 \subseteq C$ egy konvex halmaz, valamint $f : C \to \mathbb{R}^n$ függvény, amely folytonosan differenciálható C_0-on és teljesüljenek az alábbi feltételek:*

a) $\|Df(x) - Df(y)\| \leqslant \gamma\|x-y\| \quad (\forall x, y \in C_0)$,

b) $\|Df(x_0)^{-1}f(x_0)\| \leqslant \alpha$,

c) $\|Df(x)^{-1}\| \leqslant \beta \quad \forall x \in C_0$.

Legyenek továbbá:

$$h := \alpha\beta\gamma,$$
$$r_{1,2} := \alpha\frac{1 \mp \sqrt{1-2h}}{h}.$$

Ha $h \leqslant \dfrac{1}{2}$ és $\overline{S_{r_1}(x_0)} \subset C_0$, akkor az

$$x_{k+1} := x_k - Df(x_k)^{-1}f(x_k), \quad k = 0, 1, \ldots$$

módon definiált sorozat $S_{r_1}(x_0)$-ban marad és az f egyetlen megoldásához konvergál $C_0 \cap S_{r_2}(x_0)$-on.

2.1.8. Megjegyzés. Többdimenziós esetben a Newton-módszer alkalmazása során fontos a konvergencia globalizálása, például csillapítással, vagy a folytatás módszerével. Ugyanis a dimenzió növekedésével együtt csökkennek a gyököket körülvevő vonzási gömbök átmérői. Ekkor a konvergencia sebessége csökken, viszont nem lesz szükségünk arra, hogy „elég jó” kezdővektort adjunk meg.

2.2. A csillapított Newton-módszer algoritmusa

A (2.1.6) tétel biztosítja számunkra, hogy amennyiben elég közeli x_0 pontból indítjuk az iterációt, úgy a Newton-módszer konvergálni fog az x_* megoldáshoz. Annak érdekében, hogy ne legyen szükség ilyen feltételre a kezdővektor kiválasztásakor, globalizálhatjuk a módszert. Erre egy lehetőség a csillapított Newton-módszer alkalmazása, melyet már egydimenziós esetben láttunk az 1. fejezetben. Most következzen az általános algoritmus.
Legyenek adottak az alábbiak:

a) x_0: kezdeti vektor,

b) δx: maximális lépéstávolság (pl. annak a gömbnek a sugara, amelyben szerintünk a gyök van),

c) $maxit$: az iterációk maximális száma,

d) ε: pontosság,

e) az $f(x)$-et és a $Df(x)$-et kiszámító eljárások,

f) a főelemválasztásos LU-felbontás meghatározására szolgáló program (Ennek a programnak az outputjai: L, U, $sing$, ahol az utóbbi egy logikai változó, mely akkor igaz, ha a mátrix szinguláris).

1. $t := 1$;

2. $f_k := f(x_0)$, $nf_k := \|f_k\|_\infty$, $nf_0 := nf_k$;

3. **while** $k < maxit$ **do**

4.   **if** $nf_k \leqslant \varepsilon$ **then**

        STOP (sikeres kiszállás), output: x_k, nf_k, k;

    else

        $Df_k := Df(x_k)$, $Df_k \to [L, U, sing]$

5.   **if** $sing{=}igaz$ **then**

            STOP (sikertelen kiszállás), output: $sing$, x_k, nf_k;

        else

            $t := min(\delta x * \|Df_k\|_\infty / nf_k, t)$

6.   **if** $t < 10^{-6}$ **then**

                STOP (sikertelen kiszállás), output: x_k, nf_k, k;

            else

                $LUy := f_k \to y$, $l := 1$, $p := 0$

7.   **while** $l < 10$ **do**

                    $z := x_k - t * y$, $f_z := f(z)$, $nf_z := \|f_z\|_\infty$

8.   **if** $nf_z < nf_k$ **then**

                        $k := k + 1$, $x_{k+1} := z$, $f_k := f_z$, $nf_k := nf_z$

9.   **if** $l = 1$ **then**

                            $t := min(1, t * 1.5)$

                        end

                        $l := 10$

                    else

                        $t := t/2$, $l := l + 1$, $p := p + 1$;

                    end

                end

10.   **if** $p = 9$ **then**

                    STOP (sikertelen kiszállás), output: x_k, nf_k, k, l;

                end

            end

        end

    end

end

STOP ($k = maxit$, sikertelen kiszállás), output: x_k, nf_k, k.

Az algoritmusból való kilépést sikeresnek tekintjük, ha a 4. lépés feltétele teljesül. Ekkor ugyanis az x_k helyen a függvény normája elegendően kicsi lesz, ami azt jelenti a felhasználó számára, hogy megoldotta a nemlineáris egyenletrendszert.
Sikertelenségnek vesszük, ha:

- a Jacobi-mátrix LU-felbontását végző program szingularitást tapasztal;
- a t értéke túlságosan kicsi lesz, ugyanis ekkor nagy mértékben lassul a konvergencia;
- az l-ciklusban a t csökkentése nem hoz eredményt (nem teljesül az ún. leereszkedési kritérium);
- a k-ciklusban túl nagy az iterációk száma.

Az 5. lépésben érjük el azt, hogy a $t \leqslant \delta x * \dfrac{\|Df_k\|_\infty}{nf_k}$ teljesüljön. Ellenkező esetben ugyanis a Newton-lépés gyanusan nagy lenne, ellentmondva ezzel δx választásának.
Az l-ciklusban csökkentjük a t értékét, ha a már említett leereszkedési kritérium nem teljesül. Ezen a ponton azt használjuk ki, hogy a keresett gyök egyben az $\|f(x)\|$ minimuma is, hiszen:

$$f(x) = 0 \quad \Longleftrightarrow \quad \|f(x)\| = 0.$$

Tehát, minden x vektor, amely lokális minimumhelye $\|f(x)\|$-nak és $f(x) = 0$, az globális minimumhelye is $\|f(x)\|$-nak, valamint zérushelye $f(x)$-nek.
Amennyiben a kritérium teljesül, a lépést sikeresnek tekintjük és közelebb jutunk a gyökhöz. Ha ez próbálkozás nélkül ($l = 1$) sikerül, akkor óvatosan növeljük t-t, amíg el nem éri az 1 értéket.

2.3. A Broyden-módszer

Nemlineáris rendszerek megoldása során gyakran előfordul, hogy problémás a parciális deriváltak meghatározása. Ebben az esetben alkalmazhatjuk a Broyden-módszert, amely a szelőmódszer általánosítása. Előnye, hogy csak az $f(x_k)$ vektorokat használja a Jacobi-mátrix közelítéséhez, ezek pedig egyébként is a rendelkezésünkre állnak. Az iteráció első lépésében megadjuk a $Df(x)$ mátrixnak egy J_0 közelítését. Ezt megtehetjük például (2.2) szerint, és így megkapjuk x_1-et. Általános lépésben az alábbi képletet használja a Broyden-módszer:

$$x_{k+1} = x_k - J_k^{-1} f_k, \quad k = 1, 2, \dots,$$

,ahol $J_k := J(x_k)$ és $f_k := f(x_k)$. Ez az egydimenziós esetben a szelőmódszert eredményezi, ahol a deriváltat közelíthetjük osztott differenciával a következőképpen:

$$(J_k =) f_k' := \frac{f_k - f_{k-1}}{x_k - x_{k-1}}.$$

Ehhez hasonlót írunk fel magasabb dimenzióban is:

$$J_k := \frac{f_k - f_{k-1}}{x_k - x_{k-1}}.$$

Ebből adódik az úgynevezett kvázi-Newton-egyenlet, amely n darab feltételt ad a J_k meghatározásához:

$$J_k(x_k - x_{k-1}) = f_k - f_{k-1} =: z_k.$$

Azonban nekünk n^2 feltételre lenne szükségünk, ezért még követeljük meg, hogy J_k és J_{k-1} csupán egy egyrangú mátrixszal különbözzenek egymástól. Ez formálisan így fogalmazható meg:

$$(J_k - J_{k-1})y = 0, \quad \forall y \perp v_k, \quad v_k := x_k - x_{k-1}.$$

Alkalmas u_k-val ez a feltétel teljesül az alábbi egyenletben:

$$J_k - J_{k-1} = u_k v_k^T. \tag{2.7}$$

Az u_k vektor meghatározásához szorozzuk mindkét oldalt v_k-val, majd rendezzük át az egyelnetet:

$$(J_k - J_{k-1})v_k = u_k \|v_k\|^2 \quad \Longrightarrow \quad u_k = \frac{(J_k - J_{k-1})v_k}{\|v_k\|^2} = \frac{z_k - J_{k-1}v_k}{\|v_k\|^2}.$$

A kapott u_k-t (2.7)-be helyettesítve megkapjuk J_k-t, és így már mindent ismerünk x_{k+1} kiszámításához, amely a keresett megoldás egy újabb közelítése.

2.3.1. Megjegyzés. A gyakorlatban a módszer minden lépésében elkészítjük az éppen aktuális Jacobi-mátrix QR-felbontását, vagy felírjuk az LU-felbontást és invertálunk. Az előbbi akkor előnyös, ha a J rosszul kondícionált, különben az utóbbit alkalmazzuk különösen akkor, ha J nagyméretű ritkamátrix.

2.3.2. Megjegyzés. A Broyden-módszer konvergenciarendje $1+\dfrac{1}{2n}$. A lokális konvergenciát próbálkozhatunk csillapítással globalizálni, de nincs garancia arra, hogy elég kicsi t-re teljesülni fog a leereszkedési kritérium. Ha t többszörös csökkentése sem hoz eredményt, akkor érdemes új közelítést adni Df-re.

2.4. A folytatás módszere

A (2.1.8)-as megjegyzésben láttuk, hogy a dimenziószám növekedésével egyre fontosabb a Newton-módszer konvergenciájának globalizálása. Ezt ebben a részben a folytatás módszerével fogjuk elérni, amely a csillapítás általánosítása. Az eljárás lényege, hogy az eredeti $f(x) = 0$ rendszerünkhöz hozzáveszünk egy t paramétert úgy, hogy az alábbi két feltétel teljesüljön:

a) ha $t = 0$, akkor a megoldás ismert legyen;

b) ha $t = 1$, akkor az eredeti feladat megoldásához jussunk.

Ezt a következő függvénnyel valósítjuk meg:

$$g(x,t) := f(x) + (t-1)f(x_0) = 0, \quad x_0 \text{ adott.}$$

Így egy $n+1$ dimenziós feladatot kapunk. Vegyük észre, hogy $g(x_0, 0) = f(x_0) - f(x_0) = 0$, tehát $t = t_0 = 0$ esetén a g zérushelye x_0. Ezért a módszer első lépésében válasszuk x_0-t kezdővektornak, továbbá legyen $t = t_1 = \dfrac{1}{k}$, ahol k elegendően nagy. (Amennyiben nem konvergál a Newton-módszer, úgy növekjük k-t.) Innen megkapjuk az x_1 vektort, melynek segítségével és $t = t_2 = \dfrac{2}{k}$-val ki tudjuk számítani x_2-t, és így tovább. Az utolsó lépésben pedig $g(x_{k-1}, 1) = f(x_{k-1}) = 0$-t kapunk, ahol az x_{k-1} az x_*-nak közelítése.
Igazolható, hogy bizonyos feltételek teljesülése esetén a $g(x,t)$ függvénynek van $x(t)$ gyöke, minden elegendően kicsi $|t|$ értékre, továbbá $x(t)$ minden t-re differenciálható és eleget tesz a következő differenciálegyenletnek:

$$f'(x)\frac{dx}{dt} = -f(x_0).$$

Tehát adott $x(0) = x_0$ esetén $x(t)$ egyértelműen meghatározott. Ezzel visszavezettük a feladatot egy közönséges differenciálegyenlet kezdetiérték feladatára. Így is megoldhatjuk az egyenletrendszerünket, de most térjünk vissza a fenti algoritmusra, és fogalmazzuk meg a következőképpen:
Legyen $k \geq 1$ és $x_0 \in \mathbb{R}^n$ adott,

$m = 0, \ldots, k-1$-re az
$g(x, t_{m+1}) - g(x_m, t_m) = 0, \quad t_{m+1} := t_m + \Delta t, \quad \Delta t := \dfrac{1}{k},$

egyenletrendszer megoldásához l iterációt végzünk Newton-módszerrel, ennek eredményét x_{m+1}-gyel jelöljük, majd továbblépünk.

2.4.1. Tétel. (A folytatásos módszer elégséges konvergencia tétele) *Legyen x_* az $f(x) = 0$ rendszer gyöke és teljesüljön az egész $\mathbb{R}^n$-en, hogy*

$$\|f'(x)\| \leqslant M_1, \quad \|f'(x) - f'(y)\| \leqslant L\|x - y\|, \quad \|f'(x)^{-1}\| \leqslant \alpha.$$

Ha a folytatásos módszer k lépésszámára igaz

$$k \geq L\alpha^2\|f_0\|, \quad f_0 := f(x_0),$$

akkor minden $\varepsilon > 0$-hoz létezik olyan $l = l(\varepsilon, L, \alpha, M_1, \|f_0\|)$, a közbülső Newton-lépések száma, hogy érvényes a következő:

$$\|x_k - x_*\| \leqslant \varepsilon.$$

2.4.2. Megjegyzés. A folytatás módszerével tehát úgy érhetjük el a globális konvergenciát, hogy egy esetlegesen rosszul megválasztott x_0-ból egy olyan x_{k-1}-et állítunk elő, amely benne lesz a lokális konvergencia vonzási gömbjében, tehát jó lesz a Newton-módszer kezdővektorának.

2.4.3. Megjegyzés. Nem minden esetben célszerű használni a folytatásos módszert. Előfordulhat például, hogy túl sok idő megy el a közbülső egyenletrendszerek megoldására, vagy az f' Jacobi-mátrixa szinguáris. Ezek az esetek jelenősen megnehezíthetik a dolgunkat, viszont vannak olyan feladatok, amelyeket máshogyan nem tudunk megoldani.

2.5. Alkalmazások

A feladatok megoldásához használt programok a CD-mellékleten találhatók.

1. feladat

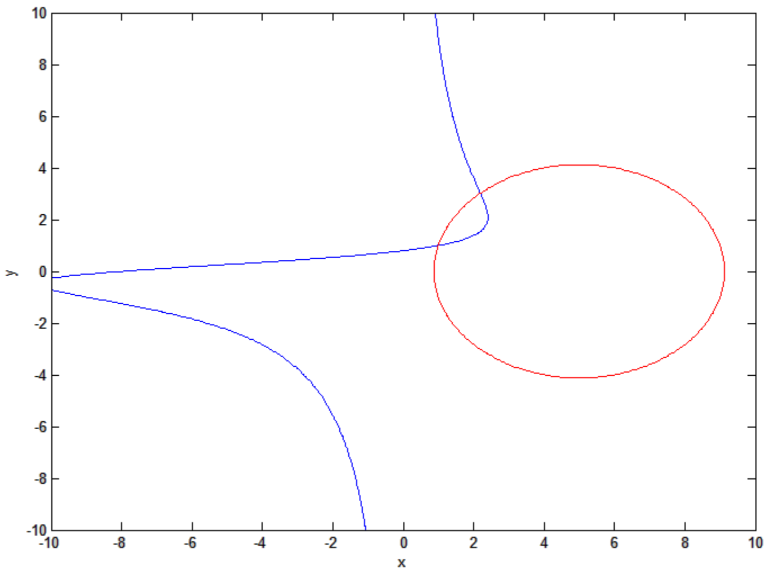
   Hány megoldása van a következő egyenletrendszernek a $[-10, 10] \times [-10, 10]$ négyzet belsejében és melyek ezek?

$$x_1^2 - 10x_1 + x_2^2 + 8 = 0 \tag{2.8}$$

$$x_1 x_2^2 + x_1 - 10x_2 + 8 = 0 \tag{2.9}$$

Megoldás:

Ábrázoljuk MATLAB segítségével az egyenleteket külön-külön implicit függvényként a megadott tartományon, és rajzoljuk ki őket egy koordinátarendszerbe a pelda1.m segítségével:



2.1. ábra. Pirossal a (2.8), kékkel a (2.9) ábrázolása

A grafikonról leolvashatjuk, hogy a $[-10, 10] \times [-10, 10]$ tartományon kettő darab gyöke van az egyenletrendszernek. Ezek megkeresésére indítsuk el különböző kezdőpontokból 'A Newton-módszer' mappában található Newton.m-et:

```
>> Newton([0.5;0.5],5e-005,@(x)[x(1).^2-10.*x(1)+x(2).^2+8 ; x(1).*x(2).^2+x(1)-10.*x(2)+8],100)
Az iterációk száma:
     3

A maradékvektor normája:
   2.4312e-05


ans =

    1.0000
    1.0000
>> Newton([3;3],5e-005,@(x)[x(1).^2-10.*x(1)+x(2).^2+8 ; x(1).*x(2).^2+x(1)-10.*x(2)+8],100)
Az iterációk száma:
     3

A maradékvektor normája:
   1.1951e-05


ans =

    2.1934
    3.0205
```

2.2. ábra. A Newton.m eredménye a $[0.5, 0.5]$ és a $[3, 3]$ kezdőpontokból

A program tehát megtalálja a két gyököt, ezek az $[1, 1]$ és a $[2.1934, 3.0205]$.

Nézzünk most egy példát arra az esetre, amikor működésbe lép a csillapított módszer t-stratégiája:

2. feladat
Oldjuk meg az alábbi rendszert a $[0, 1.4, 1]$ kezdőpontból!

$$\begin{aligned} -x_1^2 + x_3 + 3 &= 0 \\ -x_1 + 2x_2^2 - x_3^2 - 3 &= 0 \\ x_2 - 3x_3^2 + 2 &= 0 \end{aligned}$$

Megoldás:
A 'Csillapított Newton-módszer' mappából a Csillapitas.m fájl segítségével oldjuk meg az egyenletrendszert:

```
>> Csillapitas([0;1.4;1],5e-005,@(x)[-x(1).^2+x(3)+3 ; -x(1)+2.*x(2).^2-x(3).^2-3 ; x(2)-3.*x(3).^2+2],1,100)
A(z) 1. lépésben p és t értéke:
     7

    0.0078

A(z) 2. lépésben p és t értéke:
     0

    0.0117

A(z) 3. lépésben p és t értéke:
     0

    0.0176

A(z) 4. lépésben p és t értéke:
     0

    0.0264

A(z) 5. lépésben p és t értéke:
     0

    0.0396

A(z) 6. lépésben p és t értéke:
     0

    0.0593

A(z) 7. lépésben p és t értéke:
     0

    0.0890

A(z) 8. lépésben p és t értéke:
     0

    0.1335

A(z) 9. lépésben p és t értéke:
     0

    0.2002

A(z) 10. lépésben p és t értéke
     0

    0.3003

A(z) 11. lépésben p és t értéke:
     0

    0.4505

A(z) 12. lépésben p és t értéke:
     0

    0.6758

A(z) 13. lépésben p és t értéke:
     0

     1

A(z) 14. lépésben p és t értéke:
     0

     1

A(z) 15. lépésben p és t értéke:
     0

     1

Az iterációk száma:
    15

A maradékvektor normája:
   2.5572e-06

ans =

   -2.0000
    1.0000
    1.0000
```

2.3. ábra. A Csillapitas.m eredménye a $[0, 1.4, 1]$ kezdőpontból

Az eredményben szereplő p jelenti azt a számot, hogy hányszor volt szükség az egyes lépésekben a t csillapítási paraméter csökkentésére (a leereszkedési-kritérium nem teljesülése miatt). Azt látjuk, hogy az első lépésben hétszer kellett csökkenteni, így ekkor $t = 0.0078$ lett. Viszont a második lépéstől mindig nulla marad a p (elsőre teljesül a kritérium), tehát óvatosan növelhetjük t-t. A 13. lépéstől újra $t = 1$ lesz, innentől kezdve négyzetes a konvergencia. A program megtalálja a $[-2, 1, 1]$ gyököt, és a maradékvektor normája igazolja, hogy ez valóban gyök lesz.

3. feladat

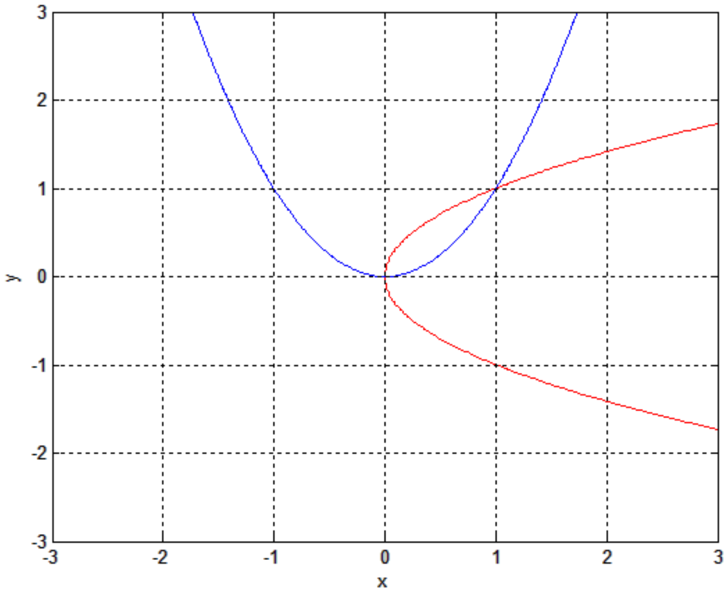
Keressük meg a következő egyenletrendszer gyökeit a $[-3, 3] \times [-3, 3]$ négyzeten és ábrázoljuk a konvergenciatartományt!

$$arctan(x_1 - x_2^2) = 0 \tag{2.10}$$

$$arctan(x_1^2 - x_2) = 0 \tag{2.11}$$

Megoldás:

Ismét ábrázoljuk MATLAB-ban implicit függvényként a két egyenletet (pelda2.m)! Ekkor az alábbi ábra adódik:



2.4. ábra. Pirossal a (2.10), kékkel a (2.11) ábrázolása

Jól látszik, hogy a $[0, 0]$ és az $[1, 1]$ lesznek a gyökök. Az első feladat mintájára indítsuk el különböző kezdőpontokból a Newton.m-et! Azt fogjuk tapasztalni, hogy a gyököknek csak egy kis környezetéből konvergál a Newton-módszer:

```
>> Newton([0.1;0.1],5e-005,@(x)[atan(x(1)-x(2)^2); atan(x(1)^2-x(2))],100)
Az iterációk száma:
     3
A maradékvektor normája:
   9.5173e-07

ans =

   1.0e-06 *

    0.6730
    0.6730
>> Newton([-1;-1],5e-005,@(x)[atan(x(1)-x(2)^2); atan(x(1)^2-x(2))],100)
Az iterációk száma:
     4
A maradékvektor normája:
   2.2386e-05
ans =

    1.0000
    1.0000
```

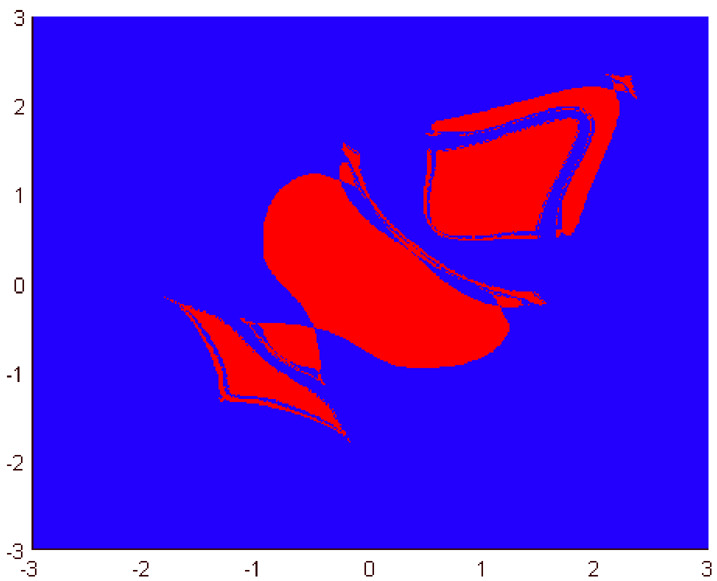
2.5. ábra. A Newton.m eredménye a $[0.1, 0.1]$ és a $[-1, -1]$ kezdőpontokból

```
>> Newton([2;1],5e-005,@(x)[atan(x(1)-x(2)^2); atan(x(1)^2-x(2))],100)
A Jacobi-mátrix szinguláris

>> Newton([10;10],5e-005,@(x)[atan(x(1)-x(2)^2); atan(x(1)^2-x(2))],100)
A Jacobi-mátrix szinguláris
```

2.6. ábra. A Newton.m eredménye a $[2, 1]$ és a $[10, 10]$ kezdőpontokból

A program megtalálja a $[0.1, 0.1]$ pontból a $[0, 0]$, a $[-1, -1]$ pontból pedig az $[1, 1]$ gyököt. Viszont a $[2, 1]$ és a $[10, 10]$ pontokból a Jacobi-mátrix szingularitására hivatkozva kilép, tehát nem konvergál. Most a 'Konvergenciatartomány ábrázolása' mappában található kirajzol.m segítségével a következő ábrát kapjuk:



2.7. ábra. Az $[arctan(x_1 - x_2^2) = 0; arctan(x_1^2 - x_2) = 0]$ egyenletrendszer konvergenciatartománya

A gafikonon piros szín jelöli azokat a pontokat, amelyekből indítva konvergált, kék szín pedig azokat, amelyekből divergált a Newton-módszer.

4. feladat
Tekintsük a 3. feladat egyenletrendszerét és indítsuk el a folytatásos módszer programját egy konvergenciatartományon kívül eső kezdőpontból!

Megoldás:

A folytatás módszere globalizálja a konvergenciát, erre látunk most egy példát. Nyissuk meg 'A folytatás módszere' mappából a Folytatas.m-et, és adjuk meg a $[10, 10]$ kezdővektort:

```
>> Folytatas([10;10],5e-005,@(x)[atan(x(1)-x(2)^2); atan(x(1)^2-x(2))],100,100)
A megtalált gyök:
    1.0000
    1.0000

Az összes iterációk száma:
   198

A maradékvektor normája:
   5.0447e-06
```

2.8. ábra. A Folytatas.m eredménye a $[10, 10]$ kezdőpontból

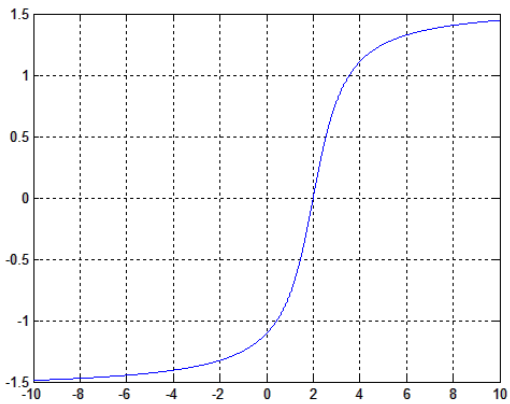
Tehát a folytatás módszere egy olyan kezdővektorból is megtalálta az $[1, 1]$ gyököt, amelyből a Newton-módszer nem konvergált.

5. feladat

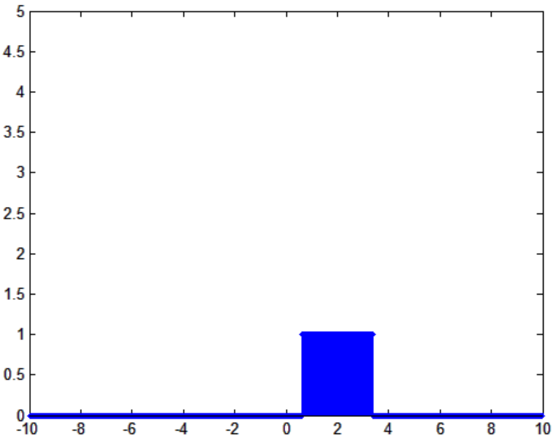
Ábrázoljuk, hogy hogyan konvergál a folytatás módszere a kezdőpontból a gyökhöz az $arctan(x - 2) = 0$ egyenlet esetén!

Megoldás:

Először rajzoljuk ki a függvényt (pelda3.m), majd azt az intervallumot, amelyből indítva a Newton-módszer konvergál ('Konvergenciatartomány ábrázolása' $\rightarrow$ arctg.m):



2.9. ábra. Az $f(x) = arctan(x - 2)$ függvény



2.10. ábra. Az $f(x) = arctan(x-2)$ függvény konvergenciaintervalluma

Az első ábráról leolvashatjuk, hogy az $x = 2$ lesz a függvény zérushelye. A második ábrán nulla a függvényérték azokon a helyeken, amelyekből nem konvergált, egy pedig azokon, amelyekből konvergált a Newton-módszer. Indítsuk el a folytatásos módszert például az $x = 10$ pontból:

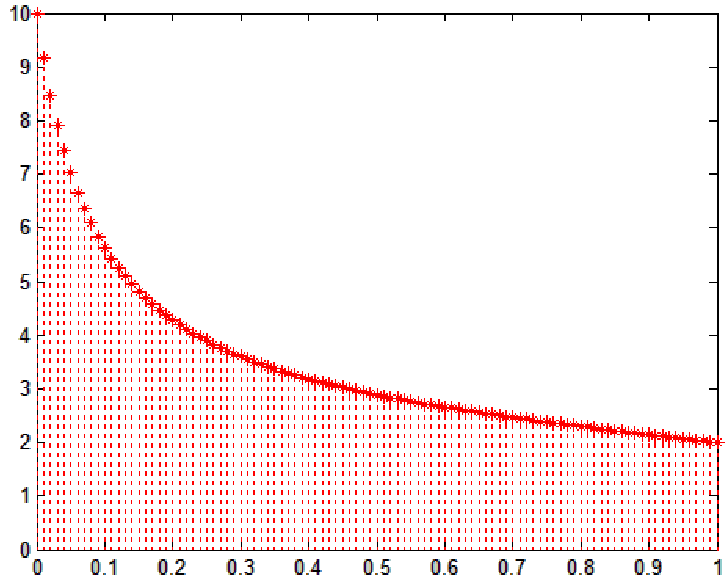
```
>> Folytatas(10,5e-005,@(x)[atan(x-2)],100,100)
A megtalált gyök:
    2.0000

Az összes iterációk száma:
   189

A maradékvektor normája:
   4.5849e-06
```

2.11. ábra. A Folytatas.m eredménye az $x = 10$ pontból

A folytatás módszere ismét megtalálta a gyököt egy olyan pontból, amely kívül esik a konvergenciaintervallumon. Most 'A folytatás módszere' mappából indítsuk el a folyt_gyokok.m-et! Ekkor megkapjuk, hogy hogyan jutott el a módszer a 10 pontból a 2-be:



2.12. ábra. A zérushelyek változása

Ez az ábra a folytatás módszerénél definiált $g(x,t)$ függvény zérushelyeit mutatja meg a $t = 0, \frac{1}{k}, \ldots, \frac{k}{k}$ helyeken.

3. fejezet

Útmutatás a programok használatához

A CD-mellékleten öt mappát, három M-fájlt és egy text dokumentumot talál az Olvasó. Az M-fájlok a fenti feladatok implicit függvényeit, valamint az $f(x) = arctan(x-2)$ függvényt rajzolják ki. A text dokumentumban találhatók az egyes módszerek utasításai a példafeladatokra.
A mappákban a módszerek M-fájljai vannak az összes 'segédprogrammal' együtt. Például a Newton-módszer mappájában benne van a Jacobi-mátrixot kiszámító fájl is, mivel ezt használja az iteráció. A programok lefutása érdekében nagyon fontos, hogy mindig az a mappa legyen megnyitva a MATLAB-ban, amelyik eljárást éppen használni szeretnénk. Ekkor már csak az utasítást kell megadnunk a text dokumentum mintájára. Az input paraméterek száma és jelentése módszerenként eltérhet, ezekhez a programok kódjában leírás található. A folytatás módszerének és a konvergenciatartomány ábrázolásának mappájában további szkriptek találhatók, ezek is a fenti feladatokhoz tartoznak. Megemlítem még, hogy a Converge.m egy $0-1$ mátrixot eredményez, amely a pontok nem konvergál-konvergál tulajdonságát jelentik. Ezt a mátrixot a *pcolor* paranccsal rajzoltathatjuk ki a kirajzol.m segítségével.

Irodalomjegyzék

[1] Gergó Lajos, *Numerikus módszerek*, ELTE Eötvös Kiadó, 2010, 157-178

[2] J. Stoer, R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag New York Inc., 1980, 244-270

[3] Stoyan Gisbert, Takó Galina, *Numerikus módszerek I.*, ELTE-Typotex Kiadó, 1993, 322-363

[4] Stoyan Gisbert, *Numerikus matematika*, Typotex kiadó, 2007, 133-158